CAR-TR-163
CS-TR-1580

DAAK70-83-K-0018
November 1985

TRANSFORMATION OF OPTICAL FLOW BY CAMERA ROTATION

Ken-ichi Kanatani*
Center for Automation Research
University of Maryland
College Park, MD 20742

COMPUTER VISION LABORATORY

CENTER FOR AUTOMATION RESEARCH

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

CAR-TR-163
CS-TR-1580

DAAK70-83-K-0018
November 1985

TRANSFORMATION OF OPTICAL FLOW BY CAMERA ROTATION

Ken-ichi Kanatani*

Center for Automation Research
University of Maryland
College Park, MD 20742

ABSTRACT

The effect of camera rotation on the description of optical flow is analyzed. The transformation law of the parameters is explicitly given by considering infinitesimal generators and irreducible reduction of the induced representation of the 3D rotation group. The parameter space is decomposed into invariant subspaces, and the optical flow is accordingly decomposed into two parts, from which an invariant basis is deduced. A procedure is presented to test the equivalence of two optical flows and to reconstruct the necessary amount of camera rotation. The relationship with the analytical expressions for 3D recovery is also discussed.

The support of the Defense Advanced Research Projects Agency and the U.S. Army Night Vision and Electro-Optics Laboratory under Contract DAAK70-83-K-0018 (DARPA Order 3206) is gratefully acknowledged.

*Permanent Address: Department of Computer Science, Gunma University, Kiryu, Gunma 376, Japan

1. INTRODUCTION

Recovery of 3D structure and motion from a 2D image sequence is one of the most challenging problems in computer vision. Most existing schemes are classified into two types. One is the *correspondence-based approach*, which does not assume any particular model of the object except the rigidity of motion and uses point-to-point correspondence explicitly. The 3D structure and motion are recovered *numerically* [1-5]. Another is the *flow-based approach*, which employs a specific model of the object and pays attention to global characteristics of the *optical flow* such as vanishing points [6-9]. This idea is fully developed by Kanatani [10-12]; if the object is a plane, the 3D structure and motion are given *analytically* in terms of invariants with respect to coordinate changes on the image plane. These invariants are derived by means of irreducible reduction of the 2D rotation group.

Although the flow-based approach does not make use of point-to-point correspondence explicitly, the optical flow itself is usually obtained by detecting the point-to-point correspondence between two successive images, and this correspondence detection is a time consuming process [13-17]. Kanatani [18-20] proposed schemes which do not use the correspondence when the object is a planar surface. In this paper, we first summarize the analytical results of Kanatani [10-12] and then generalize Kanatani's schemes [18-20] so that those analytical results can fit in the present new setting.

2. 3D MOTION FROM FLOW PARAMETERS

We assume that the image under consideration is decomposed into planar or almost planar regions, say by the method discussed by Kanatani [10,11]. Now, attention is paid to each region regarded as planar. Take a Cartesian xy-coordinate system on the image

plane and the z-axis perpendicular to it. Let $z = px + qy + r$ be the equation of that plane. The coefficients p and q are the components of the *gradient* of the plane, and r represents the *absolute depth* from the image plane. Let $(0,0,r)$, the intersection between the plane and the z-axis, be a reference point (Fig. 1). The instantaneous rigid motion is specified by *translation velocity* (a, b, c) at the reference point and *rotation velocity* $(\omega_1,\omega_2,\omega_3)$ screwwise around it (i.e., with rotation axis orientation $(\omega_1,\omega_2,\omega_3)$ and angular velocity $\sqrt{\omega_1^2+\omega_2^2+\omega_3^3}$ (rad/sec) screwwise around it). Hence, our goal is to reconstruct the nine *structure and motion parameters* p, q, r, a, b, c, ω_1, ω_2 and ω_3 from observation of the projected image motion.

(1) PERSPECTIVE PROJECTION

Let $(0,0,-f)$, the point on the z-axis at distance f from the image plane on the negative side, be the viewpoint or *focus* of the camera. A point (X, Y, Z) in the scene is projected to $(fX/(f+Z), fY/(f+Z))$ on the image plane. If the point is on the plane $z = px + qy + r$ which is moving as described above, it is easy to show that the following *optical flow* is induced at point (x, y) on the image plane:

$$\begin{aligned} u &= u_0 + Ax + By + (Ex + Fy)x, \\ v &= v_0 + Cx + Dy + (Ex + Fy)y, \end{aligned} \tag{2.1}$$

where eight *flow parameters* are given by

$$\begin{aligned} u_0 &= \frac{fa}{f+r}, & v_0 &= \frac{fb}{f+r}, \\ A &= p\,\omega_2 - \frac{pa+c}{f+r}, & B &= q\,\omega_2 - \omega_3 - \frac{qa}{f+r}, \\ C &= -p\,\omega_1 + \omega_3 - \frac{pb}{f+r}, & D &= -q\,\omega_1 + \frac{qb+c}{f+r}, \end{aligned} \tag{2.2}$$

$$E=\frac{1}{f}(\omega_2+\frac{pc}{f+r}), \qquad F=\frac{1}{f}(-\omega_1+\frac{qc}{f+r}).$$

In other words, what we are viewing is a very restricted form of motion whose velocities are specified only by eight flow parameters u_0, v_0, A, B, C, D, E and F. If these parameters are the same, the motions seem identical to the viewer. Thus, our procedure is divided into two stages. First, we detect the flow parameters u_0, v_0, A, B, C, D, E and F from a given image sequence. Next, we compute the structure and motion parameters p, q, r, a, b, c, ω_1, ω_2 and ω_3 from these flow parameters. The second stage is performed by solving the non-linear simultaneous equations (2.2) as follows (Appendix A): First, compute

$$\begin{aligned} U_0&=u_0+iv_0, \qquad T=A+D, \qquad R=C-B, \\ S&=(A-D)+i(B+C), \qquad K=E+iF, \end{aligned} \tag{2.3}$$

where i is the imaginary unit. Hence, U_0, K and S are complex numbers. If we put $V=a+ib$, $P=p+iq$ and $W=\omega_1+i\omega_2$, then V, c, P and ω_3 are given by

$$\begin{aligned} V&=(f+r)U_0/f, \qquad c=(f+c)c', \\ P(c')&=\frac{1}{2c'}(fK-U_0/f\pm\sqrt{(fK-U_0/f)^2-4c'S}), \\ W(c')&=\frac{i}{2}(fK-U_0/f\pm\sqrt{(fK-U_0/f)^2-4c'S})+iU_0/f, \\ \omega_3(c')&=\frac{1}{2}(R+\mathrm{Re}[P(c')(W(c')^*+iU_0^*/f)]), \\ c'&=-\frac{1}{2}(T+\mathrm{Im}[P(c')(W(c')^*+iU_0^*/f)]); \end{aligned} \tag{2.4}$$

where Re[.] and Im[.] denote the real and the imaginary part respectively and * the complex conjugate. Here, P, W and ω_3 are functions of c', and c' is given by solving the last of eqns (2.4). There exists only one non-zero solution c'. In fact, if we substitute the expressions for $P(c')$ and $W(c')$ in it, the equation reduces to a cubic equation in c'

(Appendix A). Since an explicit form of the solution of a cubic equation exists, we can express the solution c' explicitly, although in in a complicated form, if we wish. However, application of an iteration scheme seems more feasible. In any case, the problem is completely solved analytically, and we find that (i) the absolute depth r is indeterminate, (ii) $a/(f+r)$, $b/(f+r)$ and $c/(f+r)$ are uniquely determined, and (iii) there exist two sets of solutions for p, q, ω_1, ω_2 and ω_3, one being true and the other spurious. However, the spurious solution disappears if two or more planar regions of the same object are observed because ω_1, ω_2 and ω_3 must be common to them. Numerical schemes of 3D recovery from point-to-point correspondence have been known [2-4] and the existence of the spurious solution was pointed out [9], but analytical expressions like eqns (2.4) have not been known.

(2) ORTHOGRAPHIC APPROXIMATION

If we take the limit $f \to \infty$ of a large focal length f in eqns (2.2), we obtain the following *orthographic approximation*:

$$\begin{gathered} u_0 = a, \qquad v_0 = b, \\ A = p\omega_2, \quad B = q\omega_2 - \omega_3, \quad C = -p\omega_1 + \omega_3, \quad D = -q\omega_1, \\ E = 0, \qquad F = 0, \end{gathered} \tag{2.5}$$

and the solution is explicitly given as follows (Appendix B):

$$\begin{aligned} V &= U_0, \qquad \omega_3 = \frac{1}{2}(R \pm \sqrt{SS^* - T^2}), \\ P &= \frac{S}{k}\exp i\left(\frac{\pi}{4} - \frac{1}{2}\arg(S) + \frac{1}{2}\arg(2\omega_3 - (R + iT))\right), \\ W &= k\exp i\left(\frac{\pi}{4} + \frac{1}{2}\arg(S) - \frac{1}{2}\arg(2\omega_3 - (R + iT))\right), \end{aligned} \tag{2.6}$$

where arg denotes the argument. Here, k is an indeterminate scale factor. Thus, (i) the

absolute depth r and the velocity c in the z-direction are indeterminate, (ii) an indeterminate scale factor k is involved, and (iii) there exist two types of solutions, one being true and the other spurious. However, the spurious solution disappears if two or more planar regions of the same object are observed because ω_1, ω_2 and ω_3 must be common to them. 3D recovery from point-to-point correspondence under orthographic projection was first studied by Ullman [2], and the fact that an indeterminate scale factor is necessarily involved was already pointed out [5]. However, analytical expressions of the solution have not been known.

(3) PSEUDO-ORTHOGRAPHIC APPROXIMATION

If we omit terms of $O(1/f^2)$ but retain terms of $O(1/f)$ in eqns (2.2), E and F are replaced by

$$E=\omega_2/f\,, \qquad F=-\omega_1/f\,, \tag{2.7}$$

respectively, which we call the *pseudo-orthographic approximation.* The solution is analytically given as follows (Appendix C):

$$V=(f+r)U_0/f\,, \qquad W=ifK\,, \qquad P=\frac{S}{fK-U_0/f}\,,$$

$$\omega_3=\frac{1}{2}(R+\mathrm{Im}[Se^{-2i\alpha}]), \qquad c=-\frac{f+r}{2}(T-\mathrm{Re}[Se^{-2i\alpha}]), \tag{2.8}$$

$$\alpha\equiv\arg(fK-U_0/f)\,.$$

Hence, (i) the absolute depth r is indeterminate, (ii) $a/(f+r)$, $b/(f+r)$ and $c/(f+r)$ are uniquely determined, and (iii) p, q, ω_1, ω_2 and ω_3 are uniquely determined. It should be noted that *no spurious solution exists.*

The parameters of eqns (2.3) have geometrical meanings [10, 11]: U_0 *translation*, T *divergence*, R *rotation*, S *shearing* and K *fanning* (Fig. 2). They are transformed by a

coordinate rotation by θ on the image plane as

$$\begin{aligned} T &\rightarrow T, \qquad R \rightarrow R, \\ U_0 &\rightarrow U_0 e^{-i\theta}, \qquad K \rightarrow K e^{-\theta}, \\ S &\rightarrow S e^{-2\theta}. \end{aligned} \tag{2.9}$$

(see Appendix D.) In other words, T and R (as well as r, c and ω_3) are (absolute) invariants of *weight* 0 (or *scalars*), U_0 and K (as well as V, P and W) are (relative) invariants of weight -1 (or *vectors*), and S is a (relative) invariant of weight -2 (or a *tensor*)[12].

3. FLOW PARAMETER ESTIMATION BY FEATURES

Let $X(x,y)$ represent the image. For example, if the image consists of gray-levels, $X(x,y)$ denotes its intensity at point (x,y). If the image consists of colors, $X(x,y)$ may be a vector valued function corresponding to R, G and B. If the image consists of points and lines, $X(x,y)$ has delta-function-like singularities. In any case, we define a *feature* of image $X(x,y)$ as a *functional*, i.e., a map $F[.]$ from the set of images $X(x,y)$ to the real numbers.

Suppose that there is an optical flow $u(x,y)$, $v(x,y)$ on the image plane and that the image is moving according to this flow. Then, if $X(x,y)$ is an image at time t, it changes at time $t-\delta t$ after a short time interval into

$$\begin{aligned} &X(x-u(x,y)\delta t, y-v(x,y)\delta t) \\ &= X(x,y) - \frac{\partial X}{\partial x} u(x,y)\delta t - \frac{\partial X}{\partial y} v(x,y)\delta t + \cdots \end{aligned} \tag{3.1}$$

Then, a feature $F[X]$ at time t changes at $t-\delta t$ into $F[X]-DF[X]\delta t + \cdots$, and the *change rate* $DF[.]$ is in general a *linear* functional in $u(x,y)$ and $v(x,y)$.

In view of the optical flow of eqns (2.1), this means that we have a *linear* equation of the form

$$DF[X]=C_1[X]u_0+C_2[X]v_0+\cdots+C_7[X]E+C_8[X]F, \tag{3.2}$$

where $C_1[.]$, ... $C_8[.]$ are functionals derived from the given feature functional $F[.]$, so that they are all known functionals. On the other hand, the change rate $DF[.]$ of feature $F[.]$ can be estimated by difference schemes. For example, observe the image at time t and compute feature $F(t)$. Next, observe the image at time $t+\delta t$ after a short time interval and compute the same feature $F(t+\delta t)$. Then, the time change $DF[X]$ is approximated by $(F(t+\delta t)-F(t))/\delta t$, or we can use a higher order numerical differentiation scheme if observations are made on three or more consecutive images. Thus, all quantities except u_0, v_0, A, B, C, D, E and F in eqn (3.2) are directly computed from an image sequence without requiring point-to-point correspondence. Since an equation of the form of eqn (3.2) provides a *linear constraint*, we obtain a set of simultaneous *linear* equations to solve for the flow parameters u_0, v_0, ..., E and F if we provide eight or more independent feature functionals $F_1[.]$, $F_2[.]$, ...

The idea of using feature functionals was suggested by Amari [21,22] and was applied to 3D recovery by Kanatani [18-20]. However, he did not divide the computation process into two stages as described here but tried to compute the structure and motion parameters p, q, r, a, b, c, ω_1, ω_2 and ω_3 directly. This leads to a set of simultaneous *non-linear* equations which are difficult to solve. He proposed an iterative scheme which traces the motion along time, starting from known initial values of p, q and r as described later. Here, however, the process is divided into two stages. We first estimate the *flow parameters* by solving a set of *linear* equations. This poses no computational problem. Then, the structure and motion parameters are computed in

analytical terms as described in the previous section.

As for the feature functionals, we can use those used by Amari [21,22] and Kanatani [18,20]. We review and modify them so that they fit in the present new setting.

(1) ANISOTROPY OF TEXTURE

Consider a surface which has a spatially homogeneous (but not necessarily isotropic) texture consisting of line segments. The 3D structure and motion are detected by checking the *anisotropy* of the texture. This method, applicable in the case of orthographic projection, was first suggested by Witkin [23] and combined with *integral geometry* or *stereology* by Kanatani [18].

Let the line texture on the image plane be dissected into infinitesimal line elements. The orientation of each line element is specified by angle θ from the x-axis. Since there are two angles for the same orientation, i.e., θ and $\theta+\pi$ designate the same orientation, we choose one of them randomly with a probability of 1/2. Let the *distribution density* $f(\theta)$ be defined in such a way that $f(\theta)d\theta$ is the summed length of those line segments, *per unit area*, whose orientations are between θ and $\theta+d\theta$. By definition, $c_0=\int_0^{2\pi} f(\theta)d\theta$ is the total length of the line segments per unit area. If the distribution is isotropic, $f(\theta)$ is constant for all θ. If the distribution is nearly isotropic, the distribution dentity $f(\theta)$ is approximated by a Fourier series up to the second order

$$f(\theta)=\frac{c_0}{2}[1+a_2\cos 2\theta+b_2\sin 2\theta],$$

$$c_0=\int_0^{2\pi} f(\theta)d\theta, \tag{3.3}$$

$$a_2=\frac{1}{c_0}\int_0^{2\pi} f(\theta)\cos 2\theta d\theta, \qquad b_2=\frac{1}{c_0}\int_0^{2\pi} f(\theta)\sin 2\theta d\theta$$

Here, first order terms do not appear because of the symmetry $f(\theta+\pi)=f(\theta)$.

If the image is changing according to orthographic optical flow (i.e., eqns (2.1) with $E=0$ and $F=0$), the Fourier coefficients c_0, a_2 and b_2 of eqns (3.3) change as follows [18,29,30]:

$$\mathrm{D}\begin{bmatrix} c_0 \\ a_2 \\ b_2 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} c_0(a_2-2) & c_0 b_2 & c_2 b_2 & -c_0(a_2+2) \\ -a_2^2+6 & -b_2(a_2-4) & -b_2(a_2^2+4) & a_2^2-6 \\ -a_2 b_2 & -b_2^2-4a_2+6 & -b_2^2+4a_2-6 & a_2 b_2 \end{bmatrix}\begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix}. \qquad (3.4)$$

Thus, c_0, a_2 and b_2 serve as feature functionals, and eqn (3.4) corresponds to eqn (3.2), although another feature must be added to determine A, B C and D uniquely.

In order to measure c_0, a_2 and b_2 from a given image, we must estimate the distribution density $f(\theta)$ from the histogram of line segment orientations. To this end, we must choose an appropriate class interval for the histogram. If it is too large, estimation becomes crude. If it is too small, the counting for each class is greatly affected by noise. This difficulty arises because the definition of the distribution density $f(\theta)$ involves *infinitesimals*, i.e., a limit taking process.

There exists a method of estimating the distribution density $f(\theta)$ which does not involve a limit taking process. This is possible by a *stereological* technique. Instead of making a histogram, we count the number of intersections between the line segments and a probe line (or equally spaced parallel scanning lines). Let $N(\theta)$ be the number of intersections per unit length of the scanning line of orientation θ. Then, the observed intersection count $N(\theta)$ is related to the distribution density $f(\theta)$ by what Kanatani [18, 30] called the (two-dimensional) *Buffon transform*:

$$N(\theta)=\int_0^{2\pi}|\sin(\theta-\theta')|f(\theta')d\theta'. \qquad (3.5)$$

If the distribution density $f(\theta)$ is given by eqns (3.3), the intersection count $N(\theta)$

becomes [18, 30]

$$N(\theta)=\frac{C_0}{2}[1+A_2\cos 2\theta+B_2\sin 2\theta],$$

$$C_0=\int_0^{2\pi} f(\theta)d\theta, \tag{3.6}$$

$$A_2=\frac{1}{C_0}\int_0^{2\pi} N(\theta)\cos 2\theta d\theta, \qquad B_2=\frac{1}{c_0}\int_0^{2\pi} N(\theta)\sin 2\theta d\theta.$$

where

$$C_0=4c_0, \qquad A_2=-\frac{1}{3}a_2, \qquad B_2=-\frac{1}{3}b_2. \tag{3.7}$$

Hence, we can use C_0, A_2 and B_2 themselves as feature functionals. They are computed by measuring the intersection count $N(\theta)$ and approximating the integrations of eqns (3.6) by appropriate summations. For example, putting $N_k=N(\pi k/N)$, $k=0,1,\ldots,N-1$, we may adopt the approximation

$$C=2\sum_{k=0}^{N-1} N_k/N,$$

$$A_2=2\sum_{k=0}^{N-1} N_k\cos\frac{2\pi k}{N}\Big/\sum_{k=0}^{N-1} N_k, \quad B_2=2\sum_{k=0}^{N-1} N_k\sin\frac{2\pi k}{N}\Big/\sum_{k=0}^{N-1} N_k. \tag{3.8}$$

Consider Fig. 3, for example. If we draw on it equally spaced parallel scanning lines whose spacing is 1/22 of one side of the square frame for orientations $\theta_k=\pi k/16$, $k=0,1,\ldots,15$ with $N=16$, i.e., at 11.25° intervals, we obtain the intersection count as shown in Fig. 4, from which we obtain $A_2=-0.172$ and $B_2=0.068$. The solid curve is the corresponding approximation of eqns (3.6). Fig. 5 is the recovered distribution density of eqns (3.3) estimated by using eqns (3.7).

From eqns (3.6) and (3.10), the change rates of C_0, A_2 and B_2 become as follows:

$$\mathrm{D}\begin{bmatrix} C_0 \\ A_2 \\ B_2 \end{bmatrix} = \frac{3}{4}\begin{bmatrix} -C_0(A_2+\frac{2}{3}) & -C_0B_2 & C_0B_2 & C_0(A_2-\frac{2}{3}) \\ A_2^2-\frac{2}{3} & B_2(A_2+\frac{4}{3}) & B_2(A_2-\frac{4}{3}) & -A_2^2+\frac{2}{3} \\ A_2B_2 & B_2^2-\frac{4}{3}A_2-\frac{2}{3} & B_2^2-\frac{4}{3}A_2-\frac{2}{3} & -A_2B_2 \end{bmatrix}\begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix}. \qquad (3.9)$$

(2) ANISOTROPY OF CONTOUR

In the above, we assumed *spatial homogeneity*, since anisotropy is expressed *per unit area*. This assumption assures that the portion of the texture newly coming into view has the same statistical characteristics as the portion of the texture going out of view. However, this assumption is not necessary if the entire planar region is viewed, i.e., if we can always identify the planar region that we are looking at. Then, the distribution density $f(\theta)$ is defined in such a way that $f(\theta)d\theta$ is just the summed length (not per unit area) of those line segments whose orientations are between θ and $\theta+d\theta$. By definition, $c_0=\int_0^{2\pi} f(\theta)d\theta$ is the total length of the line segments. If the distribution is isotropic, $f(\theta)$ is constant for all θ. If the distribution density $f(\theta)$ is approximated by the Fourier series (3.3) up to the second order, the change rates of c_0, a_2 and b_2 are given by eqn (3.4) except that the first row of the matrix is replaced by

$$c_0(a_2+2) \quad c_0b_2 \quad c_0b_2 \quad -c_0(a_2-2). \qquad (3.10)$$

If we count the number of intersections between the texture of the entire planar region in question and a probe line (or equally spaced parallel scanning lines), and if $N(\theta)$ is the number of intersections per unit length of the scanning line of orientation θ, then $N(\theta)$ and $f(\theta)$ are again related by the Buffon transform of eqn (3.5). Hence, if the distribution density is approximated by eqn (3.3), $N(\theta)$ is given by the form of eqn (3.6), and the change rates of C_0, A_2 and B_2 are given by eqn (3.9) except that the first row of the matrix is replaced by

$$-C_0(A_2-\frac{2}{3}) \quad -C_0B_2 \quad 2C_0(A_2+\frac{2}{3}). \tag{3.11}$$

An interesting application arises when the planar region has no texture but its contour is viewed. Then, the contour itself can be regarded as a texture. If the contour shape is convex, the intersection counting is equivalent to measuring the *diameter* $D(\theta)$ defined as the spacing of two parallel lines of orientation θ tangent to the contour (Fig. 6), for every line has two intersections if they exist (excluding the exceptional case of tangency). The contour shape need not be convex if the diameter is measured from outside, for in this case the *convex hull* of the contour plays the role of a texture. The convex hull is *invariant* with respect to projection; the convex hull of a projected contour is the same as the projection of the convex hull of the original contour. The diameter $D(\theta)$ and the distribution density $f(\theta)$ of the contour are related as follows [19]:

$$D(\theta)=\frac{1}{2}\int_0^{2\pi}|\sin(\theta-\theta')|f(\theta')d\theta'. \tag{3.12}$$

If this function is expressed in Fourier series as in eqn (3.6), the coefficients C_0, A_2 and B_2 change as in eqn (3.9) with the first row replaced by (3.11). Consider the two contour images C and C' of Fig. 7, for example. The diameters measured at 10° intervals of orientation are plotted in Fig. 8, where the white circles correspond to C and the black ones to C'. The solid curves are approximations of the form of eqn (3.6) with C, A_2 and B_2 computed by eqns (3.8), indicating that they fairly well characterize the data.

(3) FILTERING GRAY-LEVEL IMAGES

Suppose we are observing a sequence of gray-level images of a planar region. Amari [21, 22] suggested the use of *filtering* or *weighted averaging* for feature detection. Namely, we use

$$F[X]=\int\int_W m(x,y)X(x,y)dxdy, \tag{3.13}$$

as a feature, where $m(x,y)$ is a fixed *weight function* of the filter, and integration is done over a fixed domain or *window* W on the image plane. Suppose the area of non-zero gray-levels is localized in the window W so that $X(x,y)=0$ along the window boundary and suppose the gray-level does not depend on the gradient or the depth of the object surface. An example is letters, lying entirely in the window W, drawn on a white (or black) object surface.

If the image $X(x,y)$ changes according to eqn (3.1), the feature $F[X]$ becomes after a short time interval δt

$$\begin{aligned}&\int\int_W m(x,y)dxdy-\int\int_W m(x,y)(\frac{\partial X}{\partial x}u(x,y)+\frac{\partial X}{\partial y}v(x,y))\delta t dxdy+\cdots\\&=F[X]+\int\int_W(\frac{\partial um}{\partial x}+\frac{\partial vm}{\partial y})X\delta t dxdy+\cdots,\end{aligned} \tag{3.14}$$

where we performed integration by parts, setting integrals along the window boundary to be zero according to our assumption that $X(x,y)$ is zero at the window boundary. Thus, the change rate $DF[X]$ of the feature $F[X]$ is given by

$$DF[X]=\int\int_W(\frac{\partial u}{\partial x}m+\frac{\partial v}{\partial y}m+u\frac{\partial m}{\partial x}+v\frac{\partial m}{\partial y})Xdxdy. \tag{3.15}$$

When the optical flow is given by eqns (2.1), functionals $C_1[.]$, ..., $C_8[.]$ of eqn (3.2) become

$$\begin{aligned}&C_1[X]=\int\int_W m_x Xdxdy,\quad C_2[X]=\int\int_W m_y Xdxdy,\\&C_3[X]=\int\int_W(m+xm_x)Xdxdy,\ C_4[X]=\int\int_W ym_x Xdxdy,\\&C_5[X]=\int\int_W xm_y Xdxdy,\ C_6[X]=\int\int_W(m+ym_y)Xdxdy,\\&C_7[X]=\int\int_W(3xm+x^2m_x+xym_y)Xdxdy,\end{aligned} \tag{3.16}$$

$$C_8[X]=\iint_W (3ym+xym_x+y^2m_y)X dxdy\,,$$

where $m_x=\partial m/\partial x$ and $m_y=\partial m/\partial y$ are known functions. Thus, $C_1[.]$, ..., $C_8[.]$ can be implemented as filters. Here, we assumed that $X(x,y)=0$ at the window boundary. This assumption is not essential, and it can be removed. Instead, the expressions of the functionals $C_1[.]$, ..., $C_8[.]$ include terms of line integral along the window boundary.

(4) INTEGRATION ALONG AND INSIDE THE CONTOUR

Kanatani [20] considered the case where only the bounding contour of a planar region is observed. He proposed the use of integration along the contour C of a given fixed function $m(x,y)$,

$$F[X]=\int_C m(x,y)\,ds\,, \tag{3.17}$$

as a feature, where ds denotes the line element along the contour C. This integration is easily performed on the image by using a scheme of numerical integration [20]. Then, we see that

$$DF[X]=\int_C [u\frac{\partial m}{\partial x}+v\frac{\partial m}{\partial y}+(\frac{\partial u}{\partial x}x'^2+(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x})x'y'+\frac{\partial v}{\partial y}y'^2)m\,]ds\,, \tag{3.18}$$

where $x'=dx/ds$ and $y'=dy/ds$. When the optical flow is given by eqns (2.1), functionals $C_1[.]$, ..., $C_8[.]$ of eqn (3.2) become as follows:

$$\begin{aligned}
C_1[X]&=\int_C m_x\,ds\,, \quad C_2[X]=\int_C m_y\,ds\,,\\
C_3[X]&=\int_C [xm_x+x'^2m\,]ds\,, \quad C_4[X]=\int_C [ym_x+x'y'm\,]ds\,,\\
C_5[X]&=\int_C [xm_y+x'y'm\,]ds\,, \quad C_6[X]=\int_C [ym_y+y'^2m\,]ds\,,\\
C_7[X]&=\int_C [x^2m_x+xym_y+(2xx'^2+yx'y'+xy'^2)m\,]ds\,,\\
C_8[X]&=\int_C [xym_x+y^2m_y+(yx'^2+xx'y'+2yy'^2)m\,]ds.
\end{aligned} \tag{3.19}$$

Hence, $C_1[.]$, ..., $C_8[.]$ can be computed on the image plane by using a scheme of numerical integration.

Kanatani [9] also proposed the use of surface integration inside the planar region S

$$F[X]=\int\int_S m(x,y)dxdy, \tag{3.20}$$

of a fixed function $m(x,y)$. Now, integration is done over a moving region S, not over a fixed window W. The change rate is expressed in two ways, due to *Green's theorem*, as follows:

$$\begin{aligned} \mathrm{D}F[X] &= \int\int_S [u\frac{\partial m}{\partial x}+v\frac{\partial m}{\partial y}+(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y})m]dxdy \\ &= \int_C (uy'-vx')mds. \end{aligned} \tag{3.21}$$

When the optical flow is given by eqns (2.1), functionals $C_1[.]$, ..., $C_8[.]$ of eqn (3.2) become

$$\begin{gathered} C_1[X]=\int_C my'ds=\int\int_S m_x\,dxdy, \quad C_2[X]=-\int_C mx'ds=\int\int_S m_y\,dxdy, \\ C_3[X]=\int_C xy'mds=\int\int_S [m+xm_x]dxdy, \\ C_4[X]=\int_C yy'mds=\int\int_S ym_x\,dxdy, \\ C_5[X]=-\int_C xx'mds=\int\int_S xm_y\,dxdy, \\ C_6[X]=-\int_C yx'mds=\int\int_S [m+ym_y]dxdy, \\ C_7[X]=\int_C (x^2y'-xyx')mds=\int\int_S [3xm+x^2m_x+xym_y]dxdy, \\ C_8[X]=\int_C (xyy'-y^2x')mds=\int\int_S [3ym+xym_x+y^2m_y]dxdy. \end{gathered} \tag{3.22}$$

Hence, $C_1[.]$, ... $C_8[.]$ are computed on the image plane as either line integrals or surface integrals.

4. STEPWISE TRACING AND STEREO

According to the method described so far, the flow parameters u_0, v_0, A, B, C, D, E and F can be extracted from two (or more) consecutive images, and then the structure and motion parameters a, b, c, p, q, ω_1, ω_2 and ω_3 are determined by analytical equations. As was shown, however, there remain certain indeterminacies including the absolute depth r. These indeterminacies can be removed if a sequence of images is available and if the initial position of the surface is known [19, 20]. This becomes possible if we note the fact that if a plane $z=px+qy+r$ is moving with translation velocities a, b and c and rotation velocities ω_1, ω_2 and ω_3 as described in Section 2, the coefficients p, q and r change as

$$\frac{dp}{dt}=pq\,\omega_1-(p^2+1)\omega_2-q\,\omega_3,\quad \frac{dq}{dt}=(q^2+1)\omega_1-pq\,\omega_2+p\,\omega_3,$$
$$\frac{dr}{dt}=c-pa-qb. \tag{4.1}$$

Suppose p, q and r are known at time t. Substitution of eqns (2.2) in eqn (3.2) yields

$$\mathrm{D}F[X]=C_a[X]+C_b[X]+C_c[X]+C_{\omega_1}[X]+C_{\omega_2}[X]+C_{\omega_3}[X], \tag{4.2}$$

where $C_a[.]$, $C_b[.]$, $C_c[.]$, $C_{\omega_1}[.]$, $C_{\omega_2}[.]$ and $C_{\omega_3}[.]$ are functionals defined by

$$C_a[.]=\frac{1}{f+r}(fC_1[.]-pC_3[.]-qC_4[.]),\quad C_b[.]=\frac{1}{f+r}(fC_2[.]-pC_5[.]-qC_6[.]),$$
$$C_c[.]=-\frac{1}{f+r}(C_3[.]+C_6[.]-\frac{1}{f}(pC_7[.]+qC_8[.])),$$
$$C_{\omega_1}[.]=-(pC_5[.]+qC_6[.]+\frac{1}{f}C_8[.]),\quad C_{\omega_2}[.]=pC_3[.]+qC_5[.]+\frac{1}{f}C_7[.], \tag{4.3}$$
$$C_{\omega_3}[.]=C_5[.]-C_4[.].$$

Since p, q and r are known, $C_a[.]$, ..., $C_{\omega_3}[.]$ are known functionals. The left-hand side

of eqn (4.2), i.e., the change rate F[X] of feature $F[X]$, is obtained by a numerical differentiation scheme as described earlier. Hence, if we use six or more independent feature functionals, we obtain a set of simultaneous *linear* equations of the form of eqn (4.2) to determine a, b, c, ω_1, ω_2 and ω_3. Then, p, q and r at time $t+\delta t$ are determined by integrating eqns (4.1) by some numerical integration scheme like

$$p \leftarrow p+[pq\,\omega_1-(p^2+1)\omega_2+\omega_3]\delta t\,,\quad q \leftarrow q+[(q^2+1)\omega_1-pq\,\omega_2+p\,\omega_3]\delta t\,,$$
$$r \leftarrow r+[c-pa-qb\,]\delta t\,, \tag{4.4}$$

or some other higher order scheme. This process is repeated to determine the course of motion uniquely along time [19, 20]. During this process, small errors at each step may accumulate, so that appropriate modifications are necessary once in a while, say, by the direct method described earlier or some other source of information.

This method is also used to determine the surface orientation and position p, q and r from *stereo vision* without using point-to-point correspondence. If we move the camera by l in the negative x-direction, the object moves by l in the x-direction relative to the camera. In view of eqn (4.2), the change rate $dF[X]/dl$ of feature $F[X]$ is equal to $C_a[X]$. Similarly, $C_b[X]$ and $C_c[X]$ are directly obtained by moving the camera in the y- and the z-direction and measuring the change rate of feature $F[X]$. (In practice, of course, the camera need not be moved if the necessary number of cameras are appropriately positioned beforehand.) Then, the first three of eqns (4.3) provide a set of simultaneous equations to solve for p, q and r, since $C_a[X]$, ..., $C_{\omega_3}[X]$ are also measured on the image. First, p and q are given as a solution of

$$\begin{bmatrix} C_c[X]C_3[X]+\frac{1}{f}C_a[X]C_7[X] & C_c[X]C_4[X]+\frac{1}{f}C_a[X]C_8[X] \\ C_c[X]C_5[X]+\frac{1}{f}C_b[X]C_7[X] & C_c[X]C_6[X]+\frac{1}{f}C_b[X]C_8[X] \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} \tag{4.5}$$
$$= \begin{bmatrix} fC_c[X]C_1[X]+C_a[X](C_3[X]+C_6[X]) \\ fC_c[X]C_2[X]+C_b[X](C_3[X]+C_6[X]) \end{bmatrix},$$

and r is given by

$$r = \frac{fC_1[X]-pC_3[X]-qC_4[X]}{C_a[X]} - f = \frac{fC_2[X]-qC_5[X]-qC_6[X]}{C_b[X]} - f. \tag{4.6}$$

(If we use more than three independent feature functionals, the camera need be moved in only one direction, say, in the x-direction alone. However, this does not seem to be feasible in view of noise susceptibility.)

In the orthographic approximation $f \to \infty$, eqns (4.3) become

$$\begin{gathered} C_a[.]=C_1[.], \quad C_b[.]=C_2[.], \quad C_c[.]=0, \\ C_{\omega_1}[.]=-(pC_5[.]+qC_6[.]), \quad C_{\omega_2}[.]=pC_3[.]+qC_5[.], \\ C_{\omega_3}[.]=C_5[.]-C_4[.], \end{gathered} \tag{4.7}$$

and the process goes similarly except that c is not determined, as is obvious for orthographic projection. If the feature functionals that we use are invariant with respect to translations as in (1) and (2) of the previous section, only three such features are necessary to compute ω_1, ω_2 and ω_3, which in turn determine the *trajectory* of p and q. Fig. 10 shows the trajectory of the motion of Fig. 9 obtained by measuring the diameter $D(\theta)$ [19]. However, special care should be taken when $p=0$ and $q=0$, in which case both $C_{\omega_1}[X]$ and $C_{\omega_2}[X]$ vanish and hence ω_1 and ω_2 are not determined. In this case, we must use a higher order expression of the optical flow as shown in [18, 19].

In the pseudo-orthographic approximation, the process goes similarly except that $C_c[.]$ of eqns (4.3) is replaced by

$$C_c[.] = -\frac{1}{f+r}(C_3[.] + C_6[.]). \tag{4.8}$$

Acknowledgement. The author wants to express his special thanks to Professor Azriel Rosenfeld and Professor Larry S. Davis at the University of Maryland for helpful comments. He also wants to thank Professor Shun-ichi Amari at Tokyo University, Dr. Allen Waxman at Thinking Machines Corporation and Mr. Muralidhara Subbarao at the University of Maryland for discussions and suggestions.

REFERENCES

[1] S. Ullman, The interpretation of structure from motion, *Proc. R. Soc. Lond.*, **B-203** (1979), 405 - 426.

[2] H.-H. Nagel, Representation of moving rigid objects based on visual observations, *Computer*, **14**-8 (1981), 29 - 39.

[3] H. C. Longuet-Higgins, A computer algorithm for reconstructing a scene from two projections, *Nature*, **239** (1981), 133 - 135.

[4] R. Y. Tsai and T. S. Huang, Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces, *IEEE Trans. Pattern Anal. Machine Intell.*, **PAMI-6** (1984), 13 - 27.

[5] K. Sugihara and N. Sugie, Recovery of rigid structure from orthographically projected optical flow, *Comput. Vision Graphics Image Processing*, **27** (1984), 309 - 320.

[6] D. A. Gordon, Static and dynamic visual fields in human space perception, *J. Opt. Soc. Am.*, **55** (1965), 1296 - 1303.

[7] W. F. Clocksin, Perception of surface slant and edge labels from optical flow: a computational approach, *Perception*, **9** (1980), 253 - 269.

[8] J. J. Koenderink and A. J. van Doorn, Exterospecific component of the motion parallax field, *J. Opt. Soc. Am.*, **71** (1981), 953 - 957.

[9] H. C. Longuet-Higgins, The visual ambiguity of a moving plane, *Proc. R Soc. Lond.*, **B-223** (1984), 165 - 175.

[10] K. Kanatani, *Analysis of Structure and Motion from Optical Flow: Part I Orthographic Projection*, Technical Report, University of Maryland, 1985.

[11] K. Kanatani, *Analysis of Structure and Motion from Optical Flow: Part II Central Projection*, Technical Report, University of Maryland, 1985.

[12] K. Kanatani, *Analysis of Structure and Motion from Optical Flow: Part III Invariant Decomposition*, Technical Report, University of Maryland, 1985.

[13] S. Ullman, *The Interpretation of Visual Motion*, MIT Press, Cambridge, Mass., 1979.

[14] R. Jain, Dynamic scene analysis using pixel-based processes, *Computer*, **14**-8 (1981), 12 - 18.

[15] W. B. Thompson, Lower-level estimation and interpretation of visual motion, *Computer*, **14**-8 (1981), 20 - 28.

[16] B K. P. Horn and B. G. Schunk, Determination of optical flow, *Artif. Intell.*, **17** (1981), 185 - 203.

[17] J M Prager and M. A. Arbib, Computing optic flow: the MATCH algorithm and prediction, *Comput. Vision Graphics Image Processing*, **24** (1983), 271 - 304.

[18] K. Kanatani, Detection of surface orientation and motion from texture by a stereological technique, *Artif. Intell.*, **23** (1984), 213 - 237.

[19] K Kanatani, Tracing planar surface motion from projection without knowing the correspondence, *Comput. Vision Graphics Image Processing*, **29** (1985), 1 - 12.

[20] K Kanatani, Detecting the motion of a planar surface by line and surface integrals, *Comput Vision Graphics Image Processing*, **29** (1985), 13 - 22.

[21] S Amari, Invariant structures of signal and feature spaces in pattern recognition problems, *RAAG Memoirs*, **4** (1968), 553 - 566.

[22] S. Amari, Feature spaces which admit and detect invariant signal transformations, *Proc. 4th Int. Conf. Pattern Recog.*, 1978, pp. 452 - 456.

[23] A. P. Witkin, Recovering surface shape and orientation from texture, *Artif. Intell.*, **17** (1981), 17 - 45.

[24] M. G. Kendall and P. A. Moran, *Geometrical Probability*, Charles Griffin, London, 1963.

[25] R. T. DeHoff and F. N. Rhines, *Quantitative Microscopy*, McGraw-Hill, New York, 1968.

[26] E. E. Underwood, *Quantitative Stereology*, Addison-Wesley, Reading, Mass., 1970.

[27] L. A. Santalo, *Integral Geometry and Geometric Probability*, Addison-Wesley, Reading, Mass., 1976.

[28] E. R. Weibel, *Stereological Methods*, Vols. 1, 2, Academic Press, New York, 1979, 1980.

[29] K. Kanatani, Distribution of directional data and fabric tensors, *Int. J. Engng Sci.*, **22** (1984), 149 - 164.

[30] K. Kanatani, Stereological determination of structural anisotropy, *Int. J. Engng Sci.*, **22** (1984), 531 - 546.

APPENDIX A

If we substitute eqns (2.2) in eqns (2.3), we obtain

$$U_0=\frac{f\,(a+ib)}{f+r},$$

$$T=p\,\omega_2-q\,\omega_1-\frac{pa+qb+2c}{f+r}, \qquad R=-p\,\omega_1-q\,\omega_2+2\omega_3-\frac{pb-qa}{f+r},$$

$$S=p\,\omega_2+q\,\omega_1-\frac{pa-qb}{f+r}-i\,(q\,\omega_2-p\,\omega_1-\frac{pb+qa}{f+r}), \tag{A.1}$$

$$K=\frac{1}{f}\omega_2+\frac{cp}{f\,(f+r)}+i\,(-\frac{1}{f}\omega_1+\frac{cq}{f\,(f+r)}).$$

If we put $V=a-ib$, $P=p-iq$ and $W=\omega_1+i\,\omega_2$, these equations are rewritten as

$$U_0=\frac{fV}{f+r},$$

$$R+iT=2\omega_3-\frac{2ic}{f+r}-P\left(W^*+\frac{i}{f}U_0^*\right). \tag{A.2}$$

$$S=-iP\left(W-\frac{i}{f}U_0\right), \quad K=-\frac{i}{f}W+\frac{cP}{f(f+r)}.$$

Putting

$$c'=\frac{c}{f+r}, \quad W'=W-\frac{i}{f}U_0, \tag{A.3}$$

the above equations are further rewritten as

$$V=\frac{f+r}{f}U_0, \tag{A.4}$$

$$PW'^*=(2\omega_3-R)-i(2c'+T), \tag{A.5}$$

$$PW'=iS, \quad c'P-iW'=fK-\frac{1}{f}U_0. \tag{A.6}$$

Since V is given by eqn (A.4), the remaining equations are the equations to determine c', P, W' and ω_3.

First, we check whether $c'=0$ or not. If so, we have $W'=i(fK-U_0/f)$ from the second of eqns (A.6). Then, $P=S/(fK-U_0/f)$ from the first. We can conclude $c'=0$ if and only if these W' and P satisfy $PW'^*=(2\omega_3-R)-iT$ obtained from eqn (A.5). If this is satisfied (within a certain threshold), ω_3 is given by $\omega_3=(R+\mathrm{Re}[PW'^*])/2$.

Suppose we have already checked that c' is not zero. The first of eqns (A.6) is rewritten as $(c'P)(-iW')=c'S$. Hence, eqns (A.6) means that $c'P$ and $-iW'$ are the two roots of the quadratic equation

$$X^2-LX+c'S=0 \qquad (L\equiv fK-U_0/f) \tag{A.7}$$

Hence P and W' are given as functions of c' by

$$P(c')=\frac{1}{2c'}(L\pm\sqrt{L^2-4c'S}\,),\qquad W'(c')=\frac{i}{2}(L\pm\sqrt{L^2-4c'S}\,) \tag{A.8}$$

Then, eqn (A.5) gives ω_3 as a function of c' by

$$\omega_3=\frac{1}{2}(R+\mathrm{Re}[P(c')W'(c')^*]), \tag{A.9}$$

and the equation to determine c' is

$$c'=-\frac{1}{2}(T+\mathrm{Im}[P(c')W'(c')^*]). \tag{A.10}$$

Eqn (A.10) defines a unique equation although two sets of solutions exist for P, W' and ω_3. To see this, let X_1 and X_2 be the two roots of eqn (A.7). If we choose $P=X_1/c'$ and $W'=iX_2$, we have $\mathrm{Im}[PW'^*]=-\mathrm{Re}[X_1X_2^*]/c'$, while if we choose $P=X_2/c'$ and $W'=iX_1$, we have $\mathrm{Im}[PW'^*]=-\mathrm{Re}[X_1^*X_2]/c'$. Since $\mathrm{Re}[X_1X_2^*]=\mathrm{Re}[X_1^*X_2]$, $\mathrm{Im}[PW'^*]$ of eqn (A.10) remains the same for both cases.

If we actually substitute eqns (A.8) in eqn (A.10), we obtain

$$\sqrt{16|S|^2c'^2-8\mathrm{Re}[L^2S^*]c'+|L|^4}=-8c'^2-4Tc'+|L|^2. \tag{A.11}$$

The left-hand side is a smooth concave function (or a constant if $S=0$) passing through $(0,|L|^2)$, while the right-hand side is a smooth convex quadratic function also passing through $(0,|L|^2)$ (Fig. A). Since we know that $c'\neq0$, there exists a single unique non-zero solution c'.

If we take the squares of both sides, we obtain a cubic equation

$$c'^3-Tc'^2-\frac{1}{4}(T^2-|S|^2-|L|^2)c'+\frac{1}{8}(\mathrm{Re}[L^2S^*]-T|L|^2)=0. \tag{A.12}$$

From Fig. A, it is easy to see that this cubic equation has three real roots and that the middle one is the desired root. (The other two roots were introduced by squaring of both sides.)

APPENDIX B

Since $u_0=a$ and $v_0=b$, we only need to determine p, q, ω_1, ω_2 and ω_3. (c and r are indeterminate due to orthography.) If we substitute eqns (2.5) in eqns (2.3), we obtain

$$\begin{aligned} T &= p\,\omega_2 - q\,\omega_1, \quad R = 2\omega_3 - p\,\omega_1 - q\,\omega_2, \\ S &= p\,\omega_2 + q\,\omega_1 + i\,(q\,\omega_2 - p\,\omega_1). \end{aligned} \tag{B.1}$$

The first two equations are combined into a single equation

$$R + iT = 2\omega_3 - p\,\omega_1 - q\,\omega_2 + i\,(p\,\omega_2 - q\,\omega_1). \tag{B.2}$$

If we put $P = p - iq$ and $W = \omega_1 + i\,\omega_2$, the equations become

$$PW^* = 2\omega_3 - (R + iT), \quad PW = iS. \tag{B.3}$$

Since $|PW^*| = |PW|$, the right-hand sides must have the same modulus, i.e.

$$(2\omega_3 - (R + iT))(2\omega_3 - (R - iT)) = SS^*, \tag{B.4}$$

from which ω_3 is given by

$$\omega_3 = \frac{1}{2}(R \pm \sqrt{SS^* - T^2}). \tag{B.5}$$

From eqns (B.3), we immediately see that if W and P are a solution then so are kW and P/k where k is an arbitrary non-zero real constant. Hence, we do not lose generality if we put $W = k\exp(\arg(W))$, where k is an indeterminate scale factor. Eliminating P from eqns (B.3) by taking ratios of both sides, we obtain

$$\frac{W}{W^*} = \frac{iS}{2\omega_3 - (R + iT)} \tag{B.6}$$

Taking the argument of both sides yields

$$2\arg(W)=\frac{\pi}{2}+\arg(S)-\arg(2\omega_3-(R+iT)) \pmod{2\pi}. \tag{B.7}$$

and hence

$$\arg(W)=\frac{\pi}{4}+\frac{1}{2}\arg(S)-\frac{1}{2}\arg(2\omega_3-(R+iT)) \pmod{\pi}. \tag{B.8}$$

However, we can ignore the mod π by allowing the scale factor k to be negative. Then, W is given by the second of eqns (2.6). Finally, P is given from the second of eqns (B.3) by $P=iS/W$, and hence it is written as in eqns (2.6).

APPENDIX C

If the pseudo-orthographic approximation (2.7) is adopted, eqns (A.6) are replaced by

$$PW'=iS, \qquad W=ifK. \tag{C.1}$$

Hence, W is explicitly obtained, and $P=iS/W'=S/(fK-U_0/f)$. The remaining ω_3 and c are given from eqn (A.5) as

$$\omega_3=\frac{1}{2}(R+\mathrm{Re}[PW'^*]), \qquad c=\frac{f+r}{2}(T+\mathrm{Im}[PW'^*]). \tag{C.2}$$

If we note that

$$PW'^*=-iS\frac{fK-U_0^*/f}{fK-U_0/f}=-iSe^{-2i\alpha}, \tag{C.3}$$

we obtain eqns (2.8).

APPENDIX D

Optical flows are observed in the form of eqns (2.1) with respect to an xy-coordinate system arbitrarily fixed on the image plane. The choice of the coordinate system is completely arbitrary. Suppose we use an $x'y'$-coordinate system obtained by rotating the xy-coordinate system by angle θ counterclockwise. Then, the optical flow must bear the same form

$$\begin{aligned} u' &= u_0' + A'x' + B'y' + (E'x' + F'y')x', \\ v' &= v_0' + C'x' + D'y' + (E'x' + F'y')y', \end{aligned} \tag{D.1}$$

because we are still observing the rigid motion of a plane. In other words, the optical flow is *form invariant*. Here, the old coordinates x, y and the new coordinates x', y' are related by

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{D.2}$$

Since the velocity components are transformed as a vector, the old components u, v and the new components u', v' are also related by

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}. \tag{D.3}$$

If we substitute eqns (D.2) and (D.3) into eqns (D.1) and compare the result with eqns (2.1), we find that u_0, v_0 are transformed as a vector, A, B, C, D are transformed as a tensor, and E, F are transformed as a vector, namely,

$$\begin{bmatrix} u_0' \\ v_0' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}, \tag{D.4}$$

$$\begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \tag{D.5}$$

$$\begin{bmatrix} E' \\ F' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} E \\ F \end{bmatrix}. \tag{D.6}$$

Eqns (D.4), (D.5) and (D.6) are a linear mapping from u_0, v_0, A, B, C, D, E, F to u_0', v_0', A', B', C', D', E', F', and this mapping is a *representation*, i.e., a homomorphism, of the 2D rotation group. As is well known in group representation theory, any representation is reduced to one-dimensional *irreducible representations* due to *Schur's lemma*, since the 2D rotation group is compact and Abelian. In fact, if we define U_0, T, R and S as eqns (2.3), the above mapping is rewritten as

$$\begin{gathered} U_0' = e^{-i\theta} U_0, \quad T' = T, \quad R' = R, \\ S' = e^{-2i\theta} S, \quad K' = e^{-i\theta} K. \end{gathered} \tag{D.7}$$

As Herman Weyl pointed out, irreducible representations describe physical quantities which are inherent to the phenomenon and independent of the choice of the coordinate system. Indeed, the above parameters describe geometrical characteristics of the flow itself familiar in fluid dynamics as is stated in the text. In particular, T, R and S are obtained by resolving the matrix composed of A, B, C and D into the scalar part, the deviator (or traceless symmetric) part and the antisymmetric (or skew) part. This is not a coincidence; according to the general theorem of Weyl, *all* irreducible representations of any *tensor representation* of $SO(n)$ are obtained by a combination of these decomposition processes.

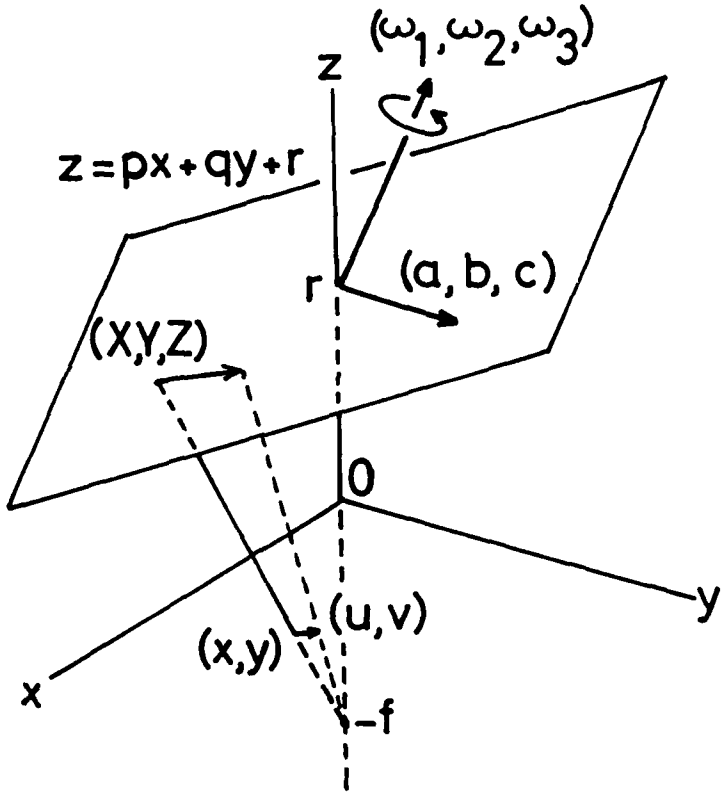

Fig. 1. A plane of equation $z = px + qy + r$ is moving with translation velocity (a, b, c) at $(0,0,r)$ and rotation velocity $(\omega_1,\omega_2,\omega_3)$ around it. An optical flow is induced on the xy-plane by perspective projection, $(0,0,-f)$ being the viewpoint.

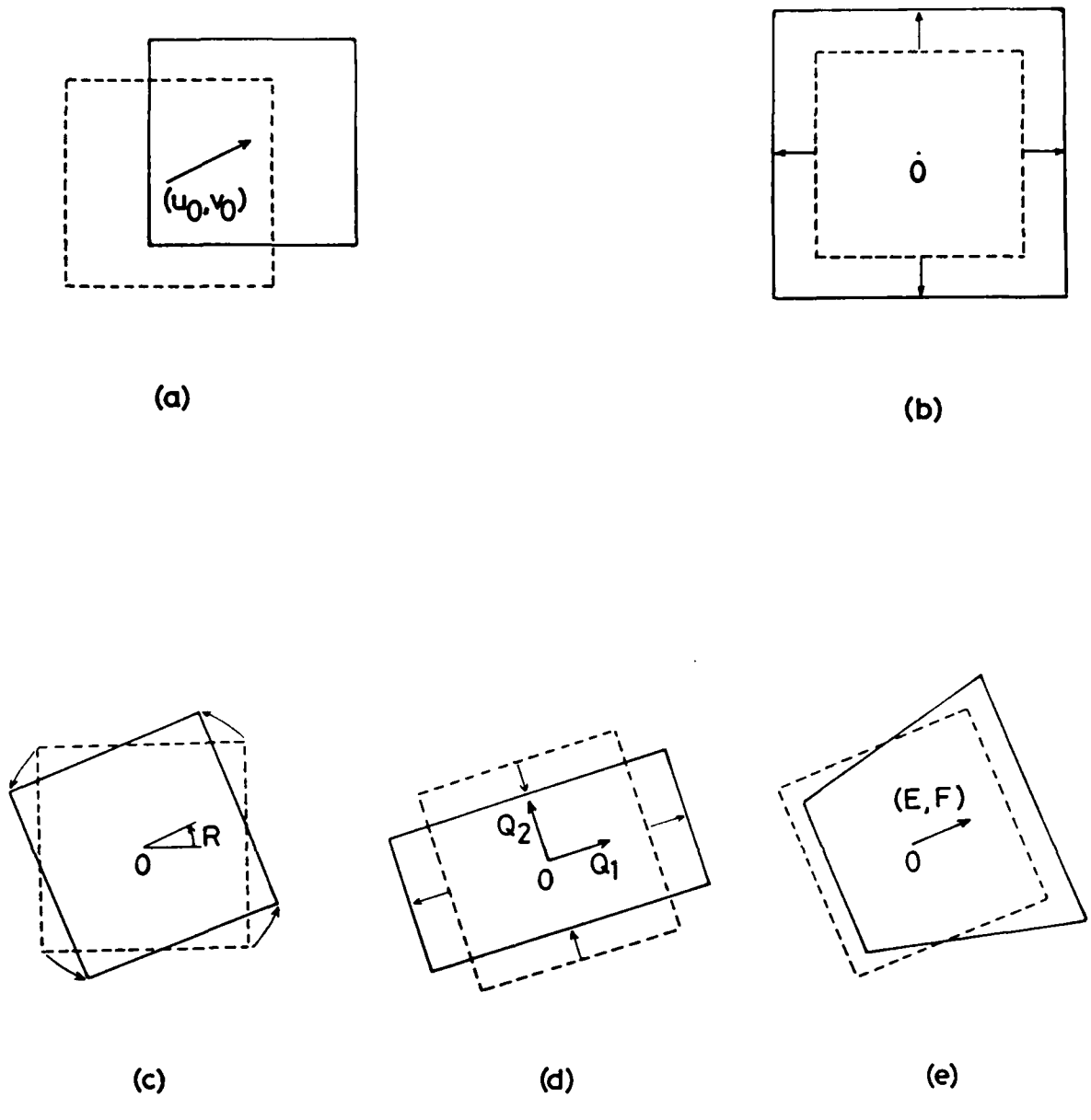

Fig. 2. (a) Translation by (u_0, v_0). (b) Divergence by T. (c) Rotation by R. (d) Shearing with $Q_1 = \exp(\arg(S)/2)$ and $Q_2 = iQ_2$ as axes of maximum extension and compression, respectively. (e) Fanning along (E, F).

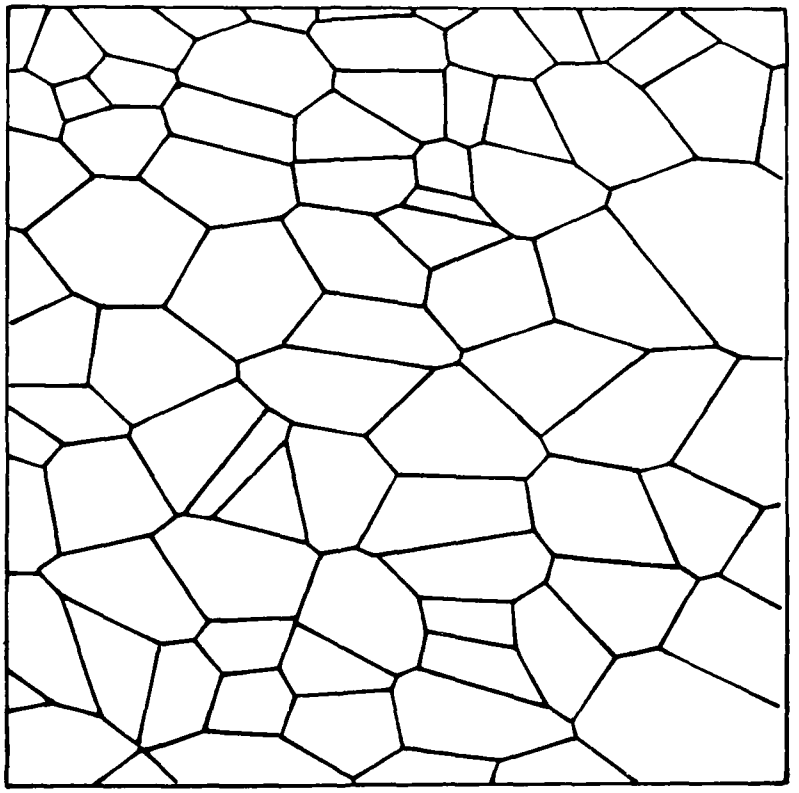

Fig. 3. An example of a textured surface image.

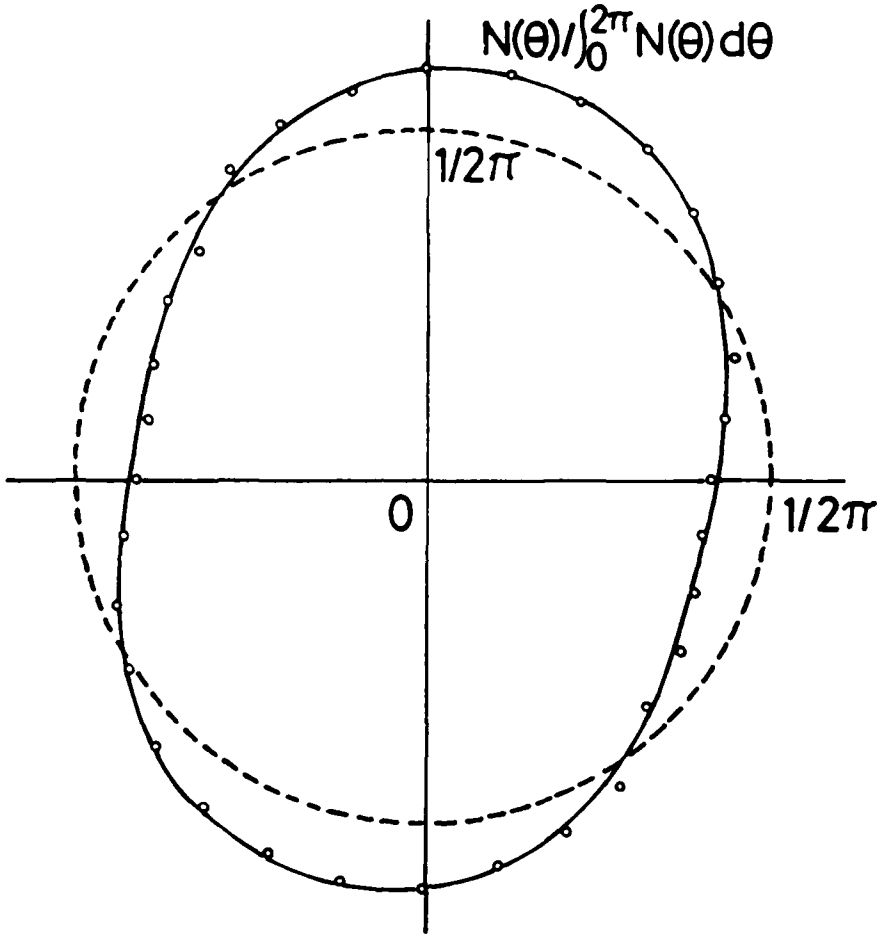

Fig. 4. The number of intersections of the texture of Fig. 1 with parallel scanning lines of different orientation, the spacing being 1/22 of the side of the square frame. The data are normalized so that the average is $1/2\pi$. The solid curve is the Fourier approximation up to the second order.

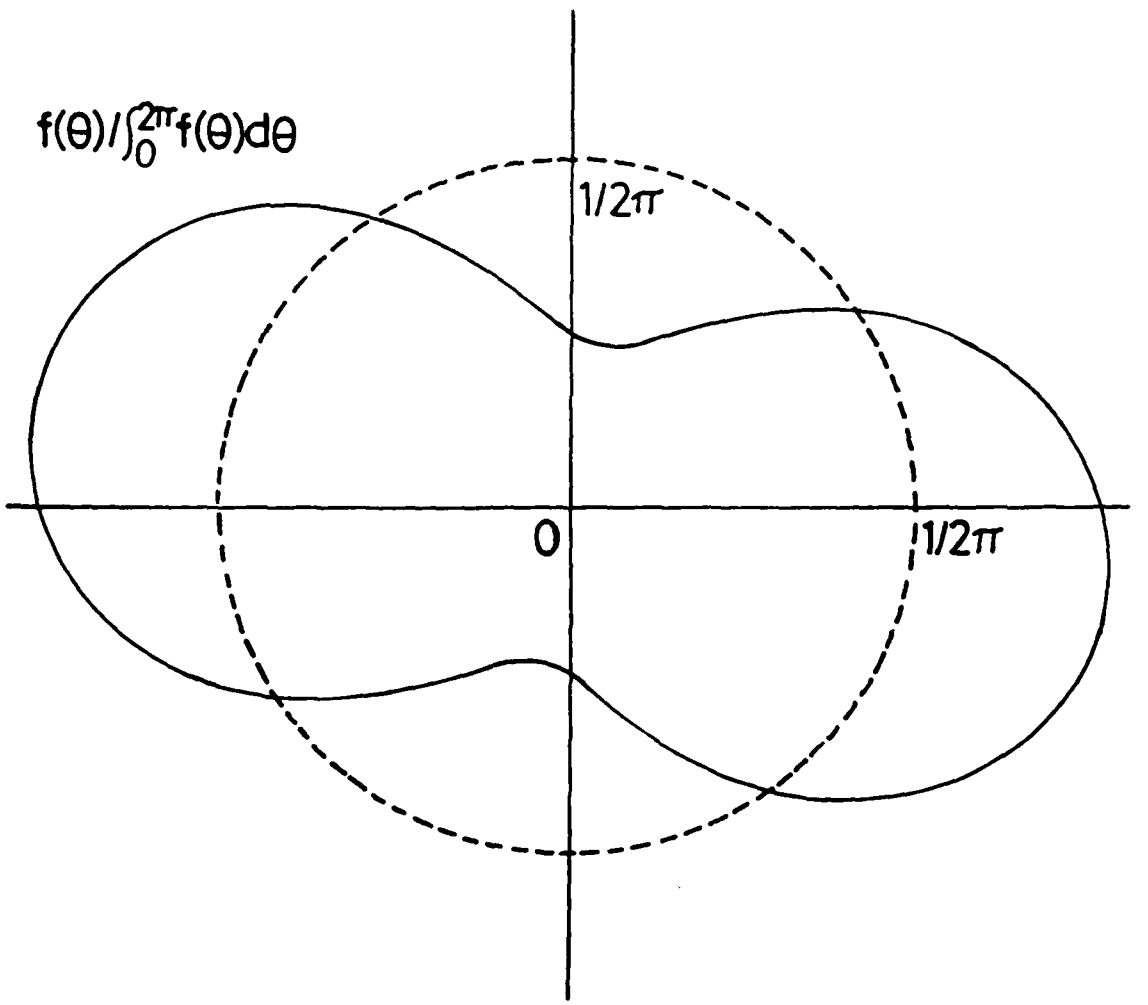

Fig. 5. Estimation of the distribution density of Fig. 3 up to the second order Fourier harmonics.

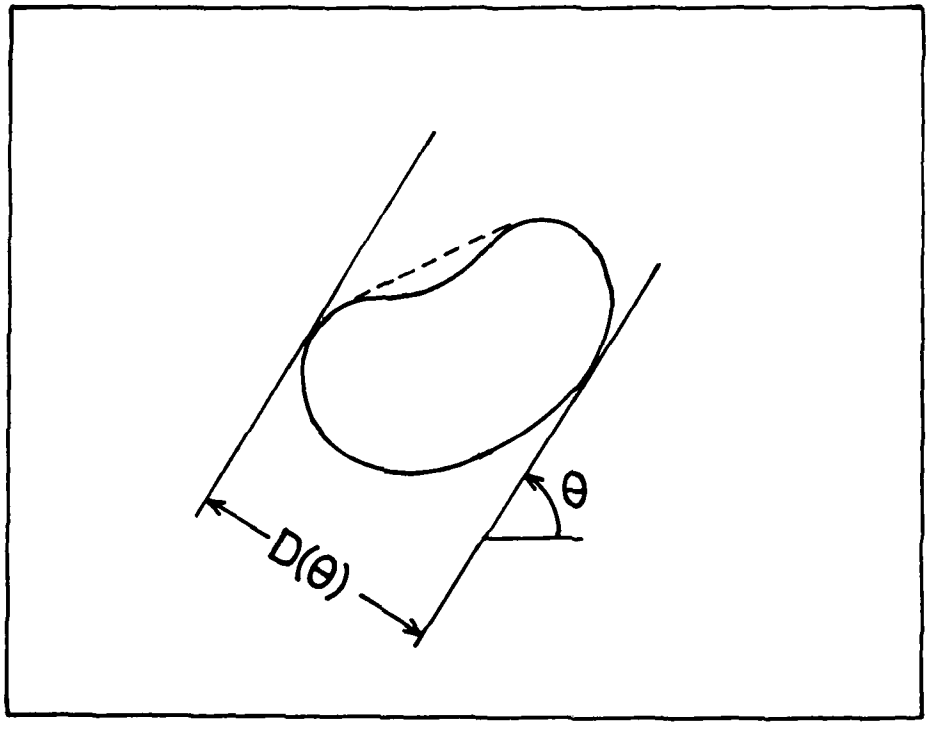

Fig. 6. The caliper diameter $D(\theta)$ is the distance between two parallel lines tangent to the contour from outside.

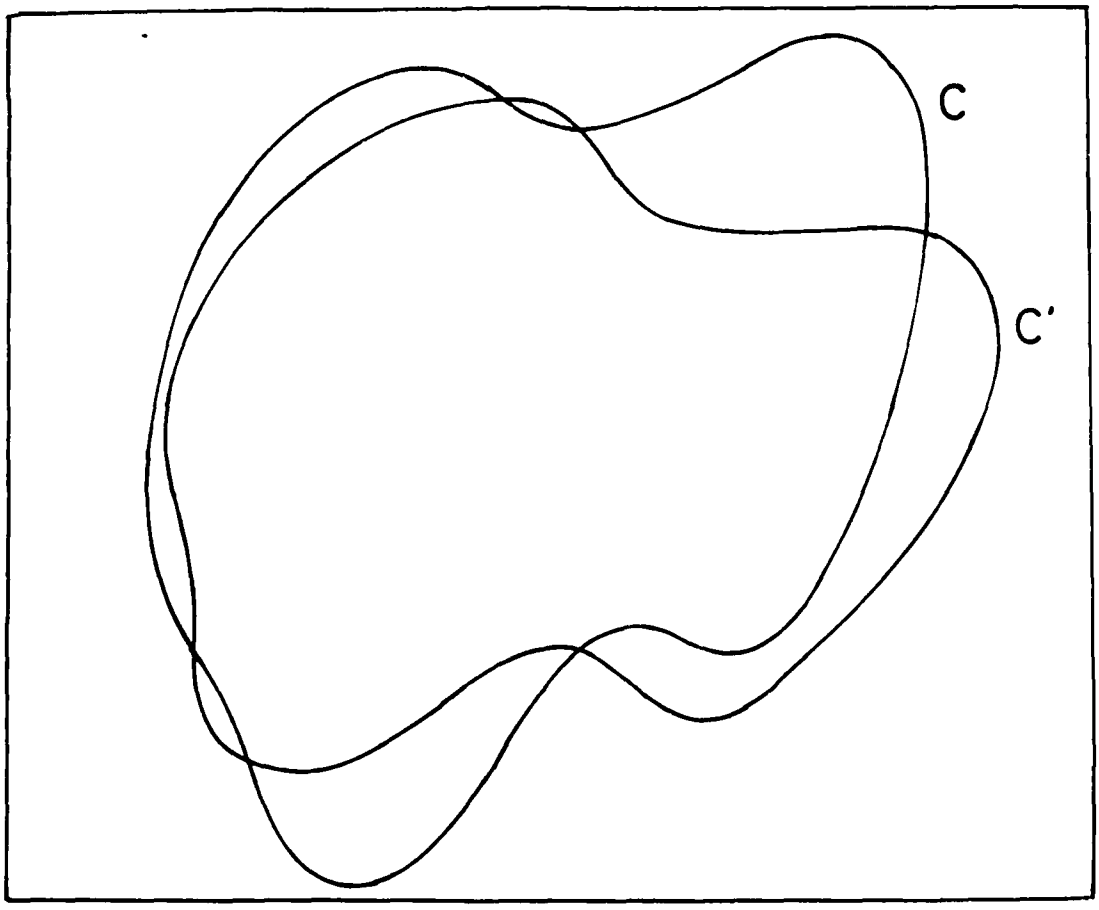

Fig. 7. Two contour images C and C' of the same planar surface.

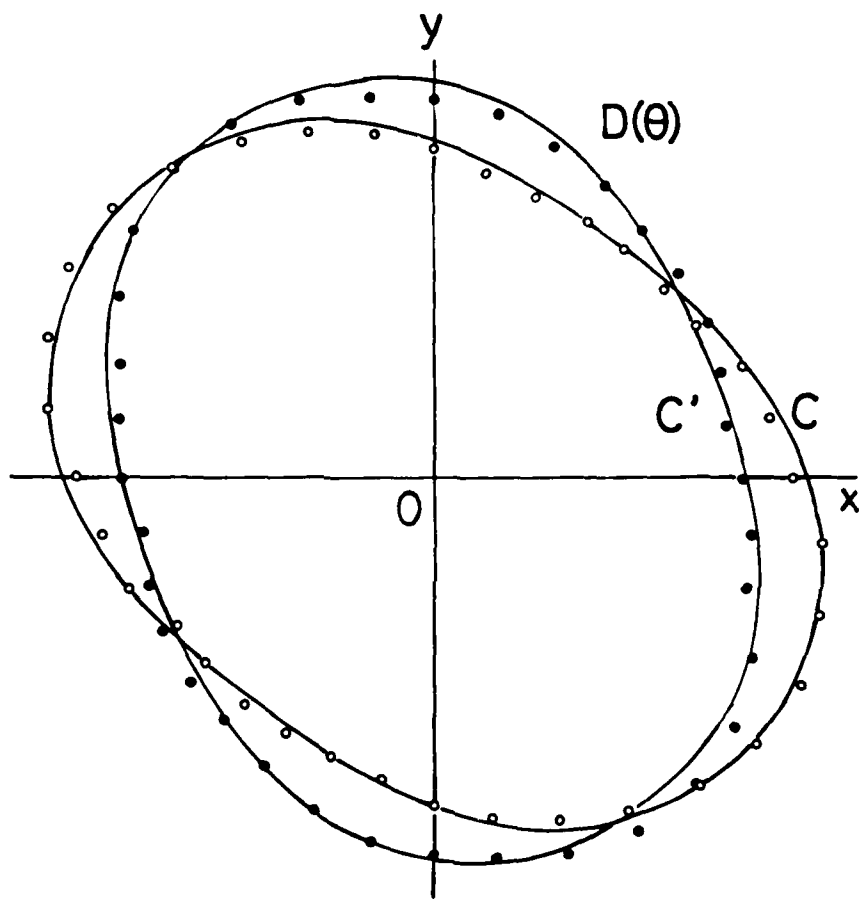

Fig. 8. Diameters of the contours C and C' of Fig. 7 for different orientations - white circles for C and black dots for C'. The solid curves are the Fourier approximation up to the second order.

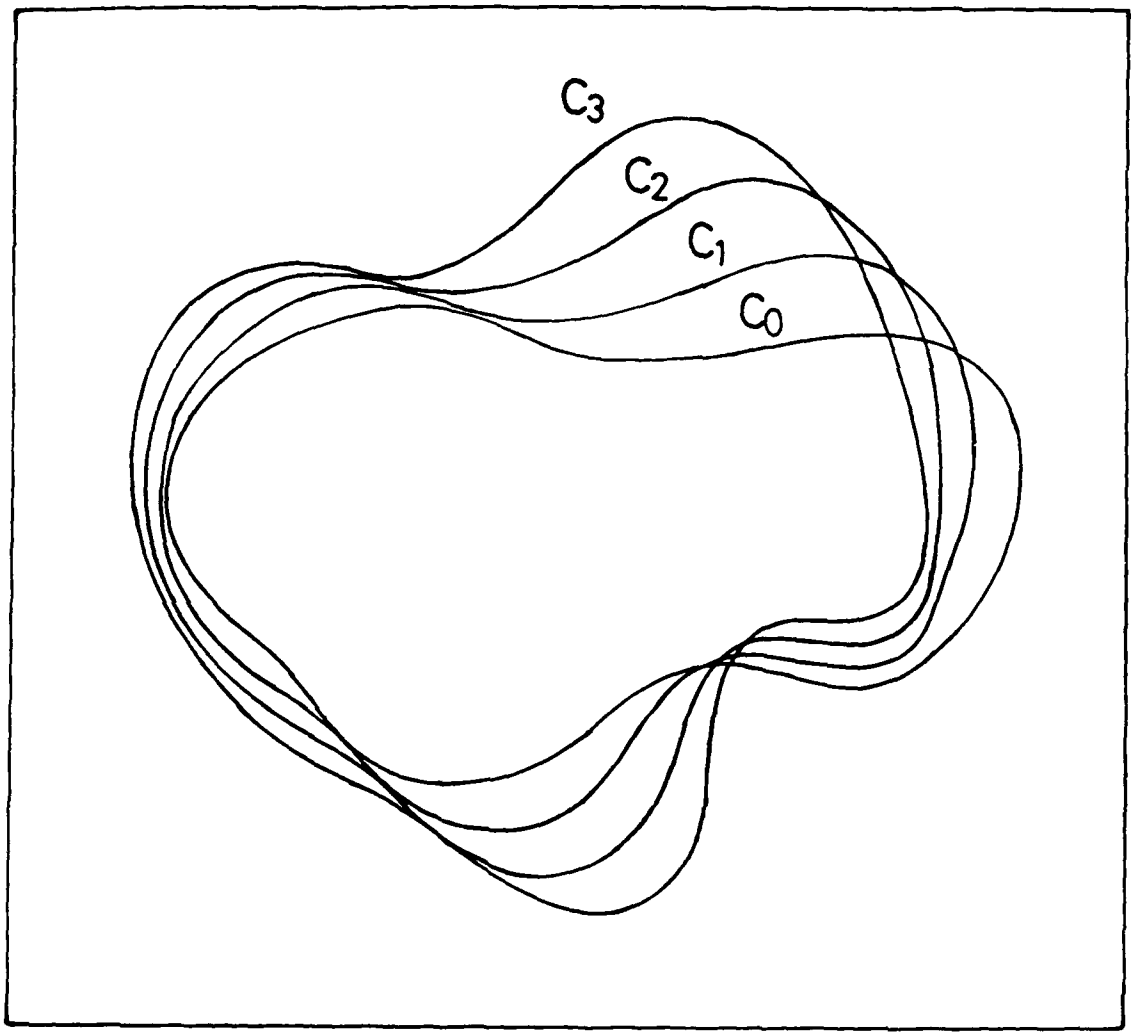

Fig. 9. Contours of a moving plane viewed orthographically. The orientation of C_0 is assumed to be known.

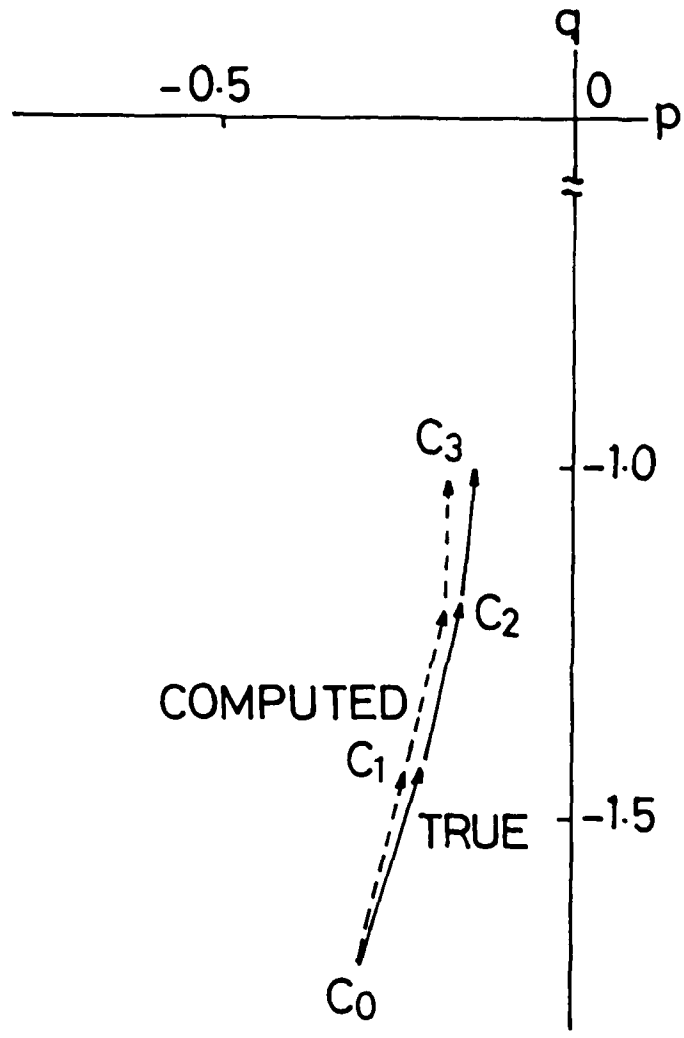

Fig. 10. The true and the computed trajectory of the gradient (p, q) obtained by measuring the diameter $D(\theta)$ of the contours of Fig. 9 at 10° intervals.

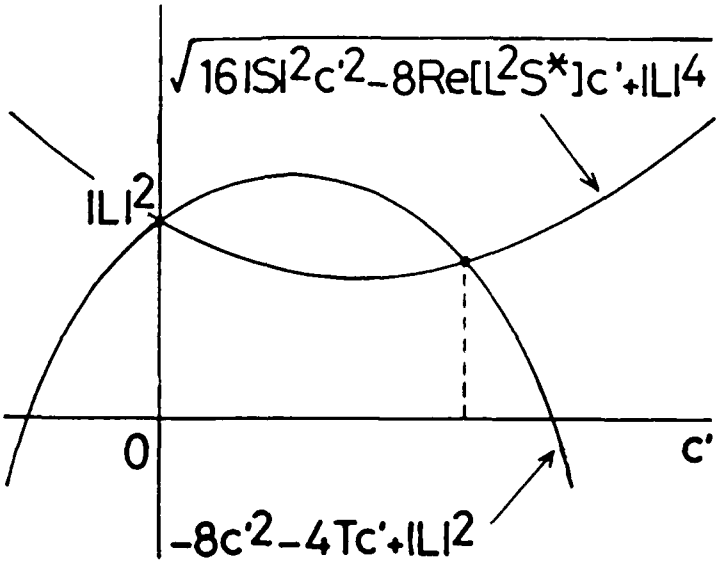

Fig. A. Existence and uniqueness of nonzero c'.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A171 611

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS N/A
2a. SECURITY CLASSIFICATION AUTHORITY N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) CAR-TR-163 CS-TR-1580	5. MONITORING ORGANIZATION REPORT NUMBER(S) N/A
6a. NAME OF PERFORMING ORGANIZATION University of Maryland 6b. OFFICE SYMBOL (If applicable) N/A	7a. NAME OF MONITORING ORGANIZATION Army Night Vision and Electro-Optics Laboratory
6c. ADDRESS (City, State and ZIP Code) Center for Automation Research College Park, MD 20742	7b. ADDRESS (City, State and ZIP Code) Fort Belvoir, VA 22060
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Defense Advanced Research Projects Agency 8b. OFFICE SYMBOL (If applicable) IPTO	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DAAK70-83-K-0018
8c. ADDRESS (City, State and ZIP Code) 1400 Wilson Vlbd. Arlington, VA 22209	10. SOURCE OF FUNDING NOS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO
11. TITLE (Include Security Classification) Transformation of Optical Flow by Camera Rotation	
12. PERSONAL AUTHOR(S) Ken-ichi Kanatani	
13a. TYPE OF REPORT Technical 13b. TIME COVERED FROM _____ TO N/A	14. DATE OF REPORT (Yr., Mo., Day) November 1985 15. PAGE COUNT 40
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The effect of camera rotation on the description of optical flow is analyzed. The transformation law of the parameters is explicitly given by considering infinitesimal generators and irreducible reduction of the induced representation of the 3D rotation group. The parameter space is decomposed into invariant subspaces, and the optical flow is accordingly decomposed into two parts, from which an invariant basis is deduced. A procedure is presented to test the equivalence of two optical flows and to reconstruct the necessary amount of camera rotation. The relationship with the analytical expressions for 3D recovery is also discussed.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code) / 22c. OFFICE SYMBOL

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE